EMC DATA DOMAIN PRODUCT OVERVIEW

Deduplication storage for next generation backup and recovery

ESSENTIALS

Scalable Deduplication Storage

- Fast, inline deduplication
- Extended retention providing up to 28.5 PB of logical storage
- 10-30x data reduction average

Easy Integration

- Supports leading backup and archive applications
- Supports leading enterprise applications for database, e-mail, content management, and virtual environments
- Simultaneous use of VTL, NAS, NDMP, and EMC Data Domain Boost

Multi-Site Disaster Recovery

- 99 percent bandwidth efficiency for network-based replication
- Flexible replication topologies for tape-free DR or tape consolidation
- Replication from up to 270 remote sites
- Encrypted replication

Ultra-Safe Storage for Reliable Recovery

- Continuous recovery verification, fault detection, and healing
- Dual disk parity RAID 6

Operational Simplicity

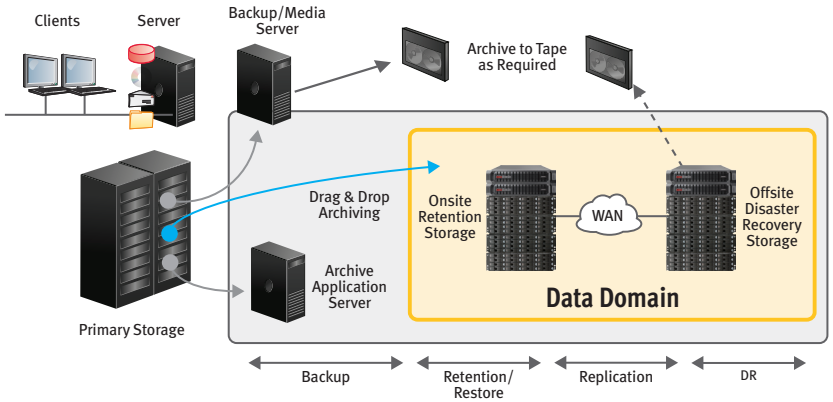
- Power, cooling, and space efficiencies for green operation
- Reduced hardware footprint
- Supports any combination of backup and archive applications in a single system

HIGH-SPEED, INLINE DEDUPLICATION AND NETWORK-EFFICIENT REPLICATION

Deduplication reduces the amount of disk storage needed to retain and protect data by ratios of 10-30x and greater, making disk a cost-effective alternative to tape. Data on disk is available online and onsite for longer retention periods, and restores become fast and reliable. Storing only unique data on disk also means that data can be cost-effectively replicated over existing networks to remote sites for disaster recovery (DR) and consolidated tape operations.

EMC® Data Domain® deduplication storage systems integrate easily with existing infrastructures and can be used seamlessly with a variety of data movers and application workloads. By consolidating to a common disk-based target, you can avoid creating disparate islands of data and storage. A single EMC Data Domain system can be used for backup and recovery, protection of enterprise applications (Oracle, Microsoft Exchange, VMware®, and others), archiving, and online reference storage.

EMC Data Domain systems deduplicate data inline during the backup process. Deduplicated data can be stored onsite for immediate restores and longer-term retention on disk. The deduplicated data can also be replicated over the WAN to a remote site for disaster recovery operations, eliminating the need for tape-based backups, or for consolidating tape backups to a central location. Data Domain systems provide flexible replication topologies to optimize your backups, such as full system mirroring, selective, bi-directional, many-to-one, one-to-many, and cascaded.

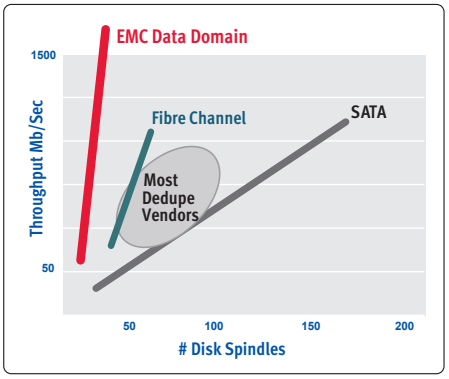


Data Domain in the Distributed Enterprise

EMC Data Domain systems deduplicate data inline during the backup process. Deduplicated data can be stored onsite for immediate restores and longer-term retention on disk. The deduplicated data can also be replicated over the WAN to a remote site for disaster recovery operations, eliminating the need for tape-based backups, or for consolidating tape backups to a central location. Data Domain systems provide flexible replication topologies to optimize your backups, such as full system mirroring, selective, bi-directional, many-to-one, one-to-many, and cascaded.

EMC DATA DOMAIN DEDUPLICATION STORAGE SYSTEMS

EMC Data Domain systems scale from smaller remote office appliances to large data center systems. These systems are available as integrated appliances or as a gateway using external storage.



CPU-Centric Storage

Data Domain Stream Informed Segment Layout (SISL) scaling architecture takes the pressure off of disk I/O as a bottleneck, so the remaining system design is CPU-centric. Other deduplication methods require more disks to increase their throughput speeds.

HIGH-SPEED, INLINE DEDUPLICATION POWERED BY CPU-CENTRIC ARCHITECTURE

Data Domain systems deduplicate data inline—during the backup process—so that the data lands on disk already deduplicated, requiring a fraction of the disk space of the original dataset. Data is "DR ready" and recoverable quickly and efficiently over existing wide area network (WAN) infrastructures.

The EMC Data Domain Stream-Informed Segment Layout (SISL™) scaling architecture leverages the continued advancement of CPU performance to add direct benefit to system throughput scalability. (For more information about SISL, see page 4.)

EASY INTEGRATION

Data Domain systems are qualified with all leading enterprise backup software and archiving applications and easily integrate into the existing storage infrastructure without change for either data center or distributed office data protection.

MULTI-SITE DISASTER RECOVERY

Data Domain systems support replication fan-in from up to 270 remote offices. Cross-site deduplication minimizes the required bandwidth between all sites, since only the first instance of data is transferred across any of the WAN segments. Datasets are effectively shrunk by 99 percent, to a size where network-efficient replication is fast and reliable.

ULTRA-SAFE STORAGE FOR RELIABLE RECOVERY

The EMC Data Domain Data Invulnerability Architecture provides continuous recovery verification along with extra levels of data protection to continuously detect and protect against data integrity issues during the initial backup and throughout the data lifecycle. Unlike any other enterprise array or file system, each appliance ensures recoverability is verified and then continuously re-verified. The systems are configured with dual disk parity RAID 6, which protects the system should two disks fail simultaneously.

OPERATIONAL SIMPLICITY

Data Domain systems are very simple to install and manage resulting in lower administrative and operational costs. All Data Domain systems have an automatic call-home system reporting capability, called autosupport, which provides email notification on complete system status. This nonintrusive alerting and data collection capability enables proactive support and service without administrator intervention, further simplifying ongoing management.

Data Domain System Specifications

	DD140	DD610	DD630	DD670[3]	DD860[3]	DD890[3]	GDA[3]	DD Archiver[3]
Logical Capacity[1,2]	9-43 TB	40-195 TB	84-420 TB	0.6-2.7 PB	1.4-7.1 PB	2.9-14.2 PB	5.7-28.5 PB	5.7-28.5 PB
Raw Capacity[2]	1.5 TB	Up to 6 TB	Up to 12 TB	Up to 76 TB	Up to 192 TB	Up to 384 TB	Up to 768 TB	Up to 768 TB
Max. Throughput (Other)	450 GB/hr	675 GB/hr[4]	1.1 TB/hr[4]	3.6 TB/hr[5]	5.1 TB/hr[5]	8.1 TB/hr[6]	10.7 TB/hr[6]	4.3 TB/hr[7]
Max. Throughput (DD Boost)	490 GB/hr	1.3 TB/hr	2.1 TB/hr	5.4 TB/hr	9.8 TB/hr	14.7 TB/hr	26.3 TB/hr	9.8 TB/hr

1. Mix of typical enterprise backup data (filesystems, databases, e-mail, developer files). The low end of capacity range represents a full backup weekly or monthly, incremental backup daily or weekly, to system capacity. The top end of the range represents full backup daily, to system capacity.
2. All capacity values are calculated using Base10 (i.e., 1TB = 1,000,000,000,000 bytes) and the maximum raw capacity configuration.
3. Includes support for add-on shelves, available separately.
4. Maximum throughput achieved using VTL interface and 4 Gbps Fibre Channel.
5. Maximum throughput achieved using Symantec OpenStorage and 10 Gb Ethernet.
6. Maximum throughput achieved using VTL interface and 8 Gbps Fibre Channel.
7. Maximum throughput achieved using CIFS and 10 Gb Ethernet.

EMC DATA DOMAIN SOFTWARE

Software options provide data protection enhancements for your environment.

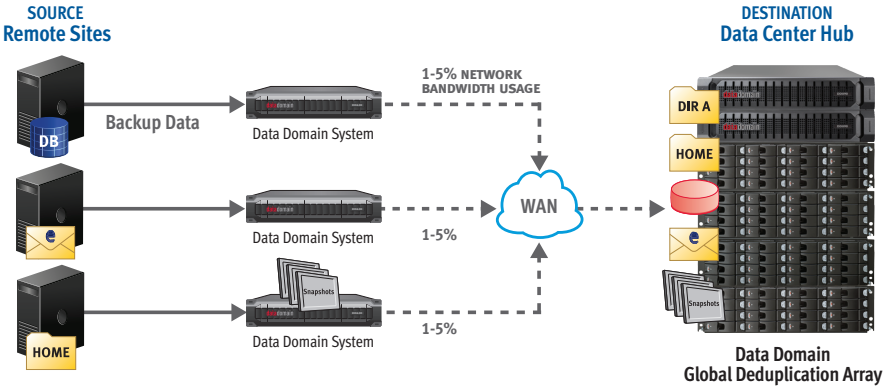
EMC Data Domain Replicator software is an automated, policy-based, network-efficient and encrypted replication software solution for disaster recovery, remote office data protection, and multi-site tape consolidation. DD Replicator software vaults (asynchronously replicates) only the compressed, deduplicated data over the WAN during the backup process, making network-based replication fast, reliable, and cost-effective. Tape backups can be eliminated or consolidated to a central site.

In order to meet a broad scope of data protection needs, DD Replicator can deploy multiple replication topologies, such as full system mirroring, one-to-many, selective, bi-directional, many-to-one and cascaded.

Two levels of bandwidth reduction are performed: local and cross-site deduplication. Data Domain deduplication massively reduces data volume stored locally, thereby reducing the amount of data that needs to be replicated. Typically less than 1 percent of a full backup, for example, is actually new, unique compressed sequences to be replicated over a WAN.

Replicator Software: 95-99 Percent Cross-Site Bandwidth Reduction

Cross-site deduplication takes place when multiple sites replicate to the same destination system. Any redundant segment previously transferred by any remote site or as a result of a local backup is not replicated again by any other remote site, saving critical network bandwidth. Cross-site deduplication improves network efficiency across all sites and reduces the storage needed at the destination, further contributing to deduplication efficiency.



EMC Data Domain Virtual Tape Library (VTL) software emulates multiple tape libraries over a Fibre Channel interface, providing deduplication storage for open systems and IBM i environments, complementing the default NAS interfaces. Data Domain VTL software eliminates tape-related failures by emulating multiple tape libraries and tape drives with up to 64,000 virtual slots across up to 250,000 virtual cartridges.

EMC Data Domain Boost software extends the optimization capabilities of Data Domain solutions. DD Boost significantly increases performance by distributing parts of the deduplication process to the backup server, and serves as a solid foundation for additional integration between backup applications and Data Domain systems.

EMC Data Domain Retention Lock software enables you to easily implement deduplication with file locking to satisfy IT governance and compliance policies for active archive protection. DD Retention Lock also enables electronic data shredding on a per-file basis to ensure that deleted files have been disposed of in an appropriate and permanent manner, in order to maintain confidentiality of classified material, limit liability, and enforce privacy requirements.

EMC Data Domain Encryption software protects backup and archive data stored on Data Domain deduplication storage systems with data encryption and compression that is performed inline— before the data is written to disk. Encrypting data at rest satisfies internal governance rules and compliance regulations and protects against theft or loss of a physical system. The combination of inline encryption and deduplication provides the most secure data-at-rest encryption solution available.

EMC DATA DOMAIN TECHNOLOGY

Data Domain technology is designed specifically to optimize the benefits of deduplication storage.

EMC Data Domain Stream-Informed Segment Layout (SISL™) scaling architecture optimizes deduplication throughput scalability and minimizes disk footprint by minimizing disk accesses. System throughput is CPU-centric, not "spindle bound." Inline deduplication throughput speeds leverage the continued advancement in CPU performance.

SISL technology first identifies 99 percent of duplicate, variable-length data segments in RAM, inline, before storing to disk. Then it stores related segments and fingerprints together, so large groups can be read at once. EMC Data Domain systems can thereby utilize the capacity of large SATA disks for data protection and—without increasing RAM—minimize the number of disks needed to deliver high throughput.

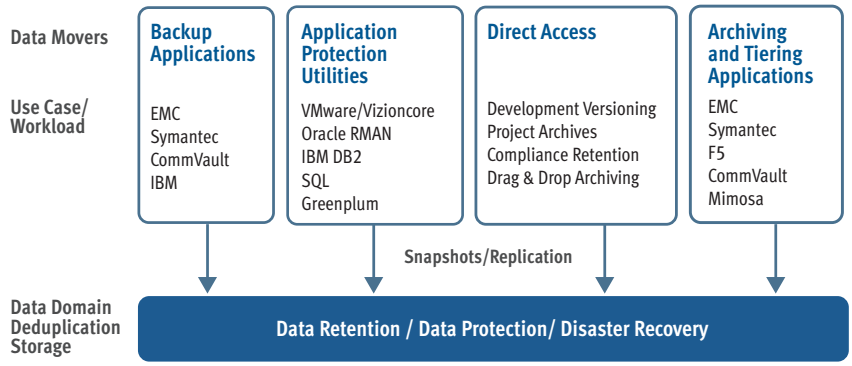
EMC Data Domain Global Compression™ technology combines high-speed, inline deduplication with local compression, after a granular, variable-length segment comparison of data. Only unique segments are written to disk. In addition, global deduplication technology and the Data Invulnerability Architecture ensure maximum data reduction and continuous recovery verification.

EMC Data Domain replication technology transfers only the deduplicated and compressed unique changes across any IP network, requiring a tiny fraction of the bandwidth, time, and cost, compared to traditional replication methods. "Time-to-DR readiness" is greatly reduced when compared to other replication methods.

EMC Data Domain Data Invulnerability Architecture offers advanced data verification and data integrity, including RAID 6 protection. Continuous fault detection, healing and write verification ensure that backup and archive data is accurately stored, available and recoverable. The no-overwrite, log-structured architecture of the EMC Data Domain filesystem together with the insistence on full-stripe writes ensures that old backups are always safe even in the face of software errors during new backups. Meanwhile, a simple and robust implementation reduces the chance of software errors in the first place.

Backup and Archive Application Support

EMC Data Domain deduplication storage systems provide a single platform for backup and archive of a broad range of enterprise applications. For data protection, all leading backup and archiving applications are supported. Users can also make Data Domain systems the target for application protection utilities like Vizioncore for virtual environments or Oracle RMAN for databases. Systems can also be accessed directly to support additional workloads. All of these data movers and workloads can be supported in the same Data Domain system, at the same time.



Backup and Archive Application Support

CONTACT US

To learn more about how EMC products, services, and solutions help solve your business and IT challenges contact your local representative or authorized reseller—or visit us at www.EMC.com

EMC Corporation
Hopkinton, Massachusetts 01748-9103
1-508-435-1000 In North America 1-866-464-7381
www.EMC.com

EMC Backup Recovery Systems
Santa Clara, California 95054
1-408-980-4800 In North America 1-866-933-3873

EMC² where information lives®